Environmental Modelling & Software 110 (2018) 107–118

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Studying the occurrence and burnt area of wildfires using zero-one-inflated structured additive beta regression

Laura Ríos-Pena [a, *], Thomas Kneib [b], Carmen Cadarso-Suárez [c], Nadja Klein [b], Manuel Marey-Pérez [d]

[a] Conservation Biology Department, CSIC Doñana Biological Station, C/Américo Vespucio 26, 41092 Seville, Spain
[b] Chair of Statistics, Georg-August-University of Göttingen, Humboldtallee 3, 37073 Göttingen, Germany
[c] Unit of Biostatistics, Department of Statistics and Operations Research, School of Medicine, University of Santiago de Compostela, C/San Francisco s/n, 15782 Santiago de Compostela, Spain
[d] PROePLA Research Group, Polytechnic Institute, University of Santiago de Compostela, Campus Universitario s/n, 27002 Lugo, Spain

ARTICLE INFO

Article history:
Received 18 January 2017
Received in revised form
13 October 2017
Accepted 15 March 2018
Available online 4 May 2018

Keywords:
Beta regression
Burnt area
Deviance information criterion
Markov chain Monte Carlo simulations
Temporal and spatial effect
Zero-one-inflated beta distribution

ABSTRACT

When studying the empirical phenomenon of wildfires, we can distinguish between the occurrence at a specific location and time and the burnt area measured. This study proposes using structured additive regression models based on zero-one-inflated beta distribution for studying wildfire occurrence and burnt area simultaneously. Beta distribution affords a convenient way of studying the percentage of burnt area in cases where such percentages are bounded away from zero and one. Inflation with zeros and ones enables observations without wildfires or with 100% burnt areas to be treated as special cases. Structured additive regression allows one to include a variety of covariates, while simultaneously exploring spatial and temporal correlations. Our inferences are based on an efficient Markov chain Monte Carlo simulation algorithm utilizing iteratively weighted least squares approximations as proposal densities. Application of the proposed methodology to a large wildfire database covering Galicia (Spain) provides essential information for improved wildfire management.

Software availability

The structured additive regression model based on the zero-one-inflated beta distribution which was used to study wildfire occurrence and burnt area simultaneously, is implemented in BayesX, which is a stand-alone software package freely available for various operating systems at http://www.bayesx.org.

All programming was performed by the authors unless stated otherwise, and the code of the model is provided in a zip file included in the supplementary material.

1. Introduction

Recent years have witnessed a rise both in the frequency of occurrence and in the burnt area of wildfires across Europe (Koutsias et al., 2015), with an increasing concentration of wildfires in the south (Urbieta et al., 2015). Though differences in climatic conditions might afford one possible explanation for this situation, this would not suffice to explain the large number of arson-related fires in this area (Viedma et al., 2015; Alcasena et al., 2015). Other potential influences include: 1) fuel load and continuity (Loepfe et al. 2014; San-Roman Sanz et al., 2013); 2) increased forest area (Catry et al., 2009); 3) the increasing wildland-urban interface (WUI) (Lampin-Maillet et al., 2010; Weise and Wotton, 2010; Paton and Tedim, 2012); 4) new patterns of agricultural land management (Verde and Zêzere, 2010; Ruiz-Mirazo et al., 2012; Guiomar et al., 2015; Viedma et al., 2015); 5) agriculture abandonment (Lasanta et al., 2015, Regos et al. 2015); 6) socio-economic changes (Curt et al., 2013; Ganteaume et al., 2013); and 7) post-fire reactions (Conedera et al., 2011). As a consequence, quantifying the risk of wildfires and managing such risk have become important issues world-wide (Thompson and Calkin, 2011).

In the past, a number of different techniques have been used to explain and predict the probability of fire occurrence, in which a binary response variable represents the presence or absence of wildfires at a given location and time point. These are: 1) linear logistic regression (LR) (González and Pukkala, 2007; Martínez-

* Corresponding author.
E-mail address: laura.rios@ebd.csic.es (L. Ríos-Pena).

https://doi.org/10.1016/j.envsoft.2018.03.008
1364-8152/

Fernández et al., 2013; Rodrigues et al., 2014); 2) generalised linear models (GLMs) (Chuvieco et al., 2010; Ordóñez et al., 2012); 3) generalised additive models (GAMs) with fixed effects, random effects and mixed effects (Brillinger et al., 2006; Reid et al., 2015); 4) spatial scan statistics (Orozco et al., 2012; Pereira et al., 2015); and 5) structured additive regression (STAR) (Ríos-Pena et al., 2015, 2017). In the above studies, different types of covariates have been used to predict the probability of wildfires, namely temperature (Huld et al., 2006; Yiannakoulias and Kielasinska, 2015), relative humidity (Padilla and Vega-García, 2011; Plucinski et al., 2014; Freeborn et al., 2015), and variables related to fuel management (González-Olabarria et al., 2012; Fréjaville and Curt, 2015).

While the probability of wildfire occurrence is certainly an interesting quantity from the perspective of risk management, it is also important to quantify the burnt area of a fire at a given location and given time point. The importance of fire can, for instance, be reflected by the percentage of burnt area (Loepfe et al., 2014; Hernandez et al., 2015; Ruffault et al., 2016). This paper proposes a new approach for simultaneously studying the occurrence and burnt area of wildfires within the framework of structured additive distributional regression, illustrated by reference to data for Galicia. The reason for using a STAR model (Fahrmeir et al., 2004) lies in the fact that non-linear covariate effects can be combined with the possibility of correcting for spatial autocorrelation through the inclusion of spatial effects. At the same time, the zero-one-inflated beta distribution incorporates the beta distribution representing the percentage of burnt area at a specific grid point. This distribution assumed for the fractional response is supplemented by zero-one-inflation. The resulting response distribution is then characterised by a location and a dispersion parameter, as well as two further parameters linked to the probabilities of zero-one-inflation (Klein et al., 2015). Each of these parameters will then be related to a regression predictor consisting of several effect types (including non-linear and spatial effects).

The main advantage of zero-one-inflated beta regression is that it enables the occurrence and burnt area of wildfires to be studied simultaneously. This therefore makes more efficient use of the information in wildfire databases that contain more than just the presence/absence of wildfires. The model also has the potential to disentangle covariate effects on the probability of effects on the burnt area occurring, and may lend better insight into the causal mechanisms underlying wildfire spread. Indeed, to date there has been no attempt to explain burnt area of wildfires by reference to fractional response regression models.

2. Dataset development

To create our own database for analysis purposes, we used the wildfire database of the Galician Regional Rural and Maritime Authority (*Consellería de Medio Rural y Mar*) (www.medioruralemar.xunta.es). This database contains information on the site (geographical coordinates) and time (hour and date) at which a fire occurs. We selected wildfires recorded during the first half of August 2006. The study period focused on an extraordinary situation in Galicia (Balsa-Barreiro and Hermosilla, 2013), inasmuch as there were a total of 2060 wildfires during the first half of August 2006 alone. The choice of this time period (15 days) was due to the fact that outside this time range the number of wildfires was quite low, which would make it almost impossible to identify the effects on the probability of occurrence or burnt area of wildfires. In addition, it should be noted that the identification of the effects in this period of time is particularly important because of the large number of wildfires that would otherwise be difficult to identify. Most of these fires were concentrated along the coastal strip and south-east of Galicia (see Fig. 1). At the same time, we obtained the percentage of area affected by wildfires, dividing Galicia into a 1×1 km cells and calculating the burnt surface area in each grid. For this, the ignition point of each wildfire has been located in one of the cells. From the burnt area data, cells have been selected totally or partially affected by the wildfire. Around 99% of the grid cells remained fire-free during the study period.

The explanatory covariates selected to predict the occurrence and burnt area of wildfires included meteorological covariates (Dickson et al., 2006) which were obtained from meteorological monitoring stations distributed throughout the study area (www.meteogalicia.es), except the wind speed covariate that can highly influence the spread of fire, but less than 10% of the meteorological monitoring stations recorded this information. Despite there being a total of 134 monitoring stations in place, only 71 were operational at the time and thus recorded data. The weather covariates selected for study purposes were daily average temperature (in °C), daily average relative humidity (in %) and rainfall (in l m^{-2}). We further restricted our analyses to meteorological stations situated at a maximum distance of 30 km from any given wildfire. Due to differences in altitude between fires and the nearest meteorological station, it was necessary to correct the temperature and rainfall covariates by reference to the difference in altimetry. The variation in daily average temperature with height was 0.65 °C/100 m to 11,000 m (Nunez and Calhoum, 1986), i.e., $tH = 15°C - 0.0065°C/m \cdot H$, where tH is the air temperature at height, and H is the difference between the wildfire and the nearest meteorological station. Likewise, the rainfall gradient amounted to an 8% variation per 100 m difference in height between any given ignition point and the meteorological station selected, i.e., $P_{(corrected)}(monthly) = P_{(monthly)} \pm 0.08$, where P is rainfall. Moreover, we studied the impact of calendar time (days), in order to capture dependence on climatic factors in the preceding days, and the altitude in each grid. The spatial component was made up of 30,158 cells, derived by dividing Galicia into a 1×1 km cells, with each cell being associated with one of the 315 municipalities in which the study area is administratively divided. The centroid of each cell is associated with a municipal area, so that each grid belongs to a municipality. The database contained a total of 452,370 records (30,158 cells per fifteen days of study). Table 1 shows all the variables used in the study, along with a brief description of each.

3. Statistical methodology

3.1. Zero-one-inflated beta regression

Zero-one-inflated beta regression is particularly well suited to the analysis of wildfire occurrence and burnt area, since it accommodates both fractional response values, i.e., responses presenting percentages, and responses observed at the extremes, i.e., clusters of observations exactly equal to zero or one. The zero-one-inflated beta distribution is defined as the mixture of one continuous component following a beta distribution and two point masses at zero and one, such that the mixed discrete continuous density is finally obtained as

$$\boldsymbol{p}(\boldsymbol{y}) = \left\{ \begin{array}{cc} \boldsymbol{\pi}_0 & \boldsymbol{y} = 0 \\ (1 - \boldsymbol{\pi}_0 - \boldsymbol{\pi}_1)\left(\dfrac{1}{\boldsymbol{B}(\boldsymbol{a}, \boldsymbol{b})}\right)\boldsymbol{y}^{\boldsymbol{a}-1}(1 - \boldsymbol{y})^{\boldsymbol{b}-1} & 0 < \boldsymbol{y} < 1 \\ \boldsymbol{\pi}_1 & \boldsymbol{y} = 1 \end{array} \right\}$$

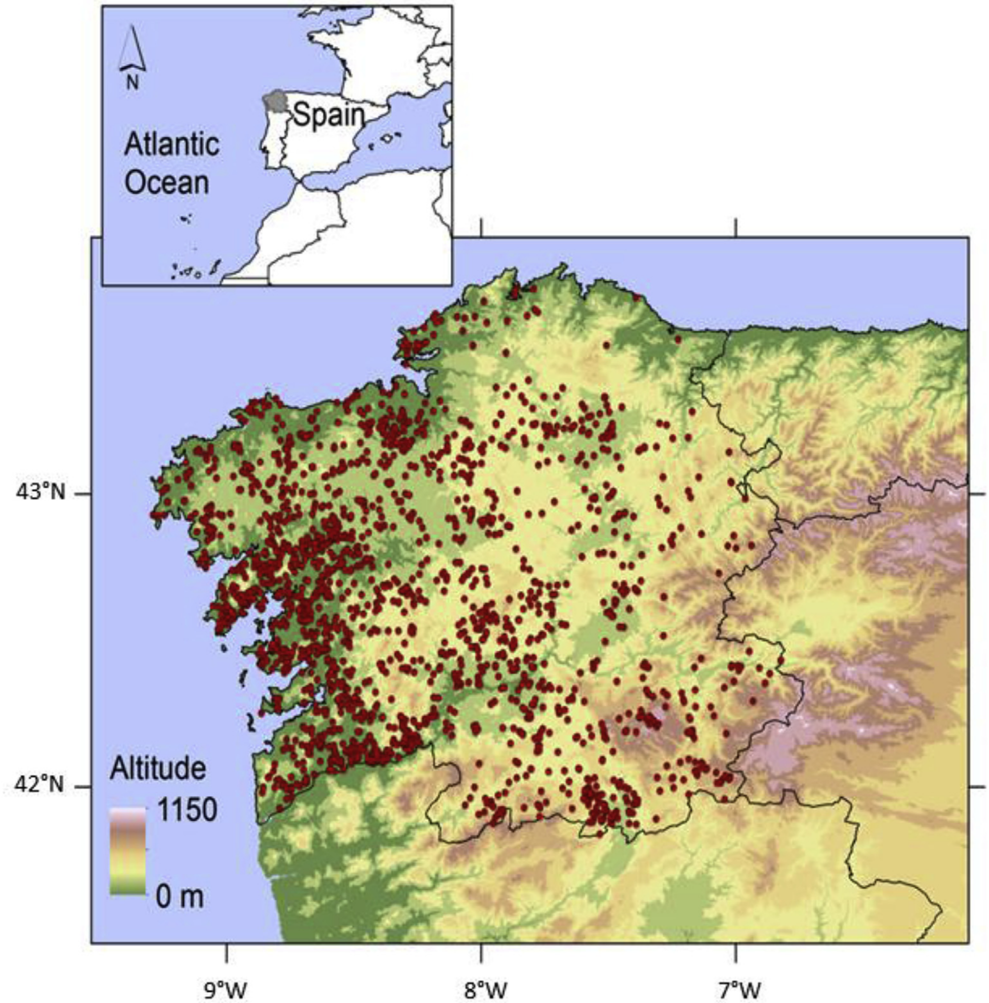

Fig. 1. Location map showing the study area (NW Spain) and spatial distribution of the wildfires that occurred during the study period. Red points mark the centroids of the wildfires. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Variables generated from baseline data.

Variable	Description
number_wildfires	Geographical coordinates of the occurrence of wildfires
burnt	Area burnt in each grid in %
altitud	Altitude of each grid in metres (m)
temp	Average daily temperature recorded for each grid in °C
hr	Daily average relative humidity for each grid in %
pp	Daily rainfall recorded for each grid in $\mathrm{l\ m^{-2}}$
day	Day of the month
cdconc	Municipalities in Galicia

where $\boldsymbol{B}(\boldsymbol{a},\boldsymbol{b})$ denotes the beta function

$$\boldsymbol{B}(\boldsymbol{a},\boldsymbol{b}) = \int_0^1 \boldsymbol{y}^{\boldsymbol{a}-1}(1-\boldsymbol{y})^{\boldsymbol{b}-1}\boldsymbol{dy}$$

with parameters $\boldsymbol{a}>0$ and $\boldsymbol{b}>0$, while $\boldsymbol{\pi}_0$ and $\boldsymbol{\pi}_1 \in (0,1)$. The probability $\boldsymbol{\pi}_0$ is the probability of no wildfire while $\boldsymbol{\pi}_1$ is the probability that the pixel is completely burnt. Therefore, $\boldsymbol{\pi}_0$ and $\boldsymbol{\pi}_1$ are the probabilities for the point masses at zero and one, respectively, such that $\boldsymbol{\pi}_0 + \boldsymbol{\pi}_1 < 1$. Ospina and Ferrari (2010) propose an alternative parameterisation of the zero-one-inflated beta distribution, which enables an easier interpretation in the regression context. Firstly, the continuous part corresponding to a beta distribution is parameterised, based on the expectation

$$\boldsymbol{\mu} = \frac{\boldsymbol{a}}{\boldsymbol{a}+\boldsymbol{b}}\,\varepsilon\,(0,1)$$

and the scale parameter

$$\boldsymbol{\sigma} = \frac{1}{\boldsymbol{a}+\boldsymbol{b}+1}\,\varepsilon\,(0,1).$$

In a second step, the probabilities for the point masses at zero and one are replaced by $\boldsymbol{\nu}>0$ and $\boldsymbol{\tau}>0$, determined from the equations

$$\boldsymbol{\pi}_0 = \frac{\boldsymbol{\nu}}{1+\boldsymbol{\nu}+\boldsymbol{\tau}} \qquad \boldsymbol{\pi}_1 = \frac{\boldsymbol{\tau}}{1+\boldsymbol{\nu}+\boldsymbol{\tau}}$$

i.e., $\boldsymbol{\nu}$ and $\boldsymbol{\tau}$ are proportional to the probabilities for no and complete damage at a given observation point, respectively. The advantage of this parameterisation is that it avoids the constraint $\boldsymbol{\pi}_0 + \boldsymbol{\pi}_1 < 1$.

Following the idea of distributional regression (Klein et al., 2015) and generalised additive models for location, scale and shape (Rigby and Stasinopoulos, 2005), we relate each of the four parameters of the zero-one-inflated beta distribution to regression

predictors, such that

$$\mu_i = \frac{\exp\left(\eta_i^{\mu}\right)}{1+\exp\left(\eta_i^{\mu}\right)}, \sigma_i = \frac{\exp\left(\eta_i^{\sigma}\right)}{1+\exp\left(\eta_i^{\sigma}\right)}, \nu_i = \exp\left(\eta_i^{\nu}\right), \tau_i = \exp\left(\eta_i^{\tau}\right).$$

Each predictor is chosen to be of a structured additive type, as explained in the following section.

3.2. Structured additive regression

Structured additive regression (Fahrmeir et al., 2004) provides a generalisation of generalised linear and additive models, where a more flexible structure is assumed for regression predictors. In generic notation, a structured additive predictor is given by

$$\eta = \beta_0 + f_1(x) + \ldots + f_J(x) \tag{1}$$

where the parameter index has been dropped for the sake of simplicity (which also suppresses the potential dependence of the number of effects J and the chosen covariates x on the parameter), β_0 is the overall intercept, and the functions $f_1, \ldots, f_J$ represent regression effects of different types of the generic covariate vector x.

Each of the functions in the additive predictor is now approximated by a linear combination of basis functions $B_1(x), \ldots, B_K(x)$, such that

$$f(x) = \sum_{k=1}^{K} \beta_k B_k(x),$$

where β_k denotes the corresponding basis coefficients. Examples include penalised splines, Markov random fields or random effects, as will be discussed in more detail in the next section. In matrix notation, a structured additive predictor of the form (1) can then be expressed as

$$\eta = \beta_0 1 + Z_1\beta_1 + \ldots + Z_J\beta_J,$$

where η is obtained by stacking all individual observations, 1 refers to a vector of ones, Z_j arises from evaluating the basis functions at the observed covariates, and β_j is the vector of corresponding basis coefficients.

To ensure favourable properties of the function estimates, such as smoothness in the case of non-linear and spatial effects, or shrinkage in the case of random effects, we rely on a Bayesian formulation where prior distributions are assigned to the vectors of regression coefficients. Generically these are given by multivariate normal distributions,

$$p\left(\beta_j | \tau_j^2\right) \propto \exp\left(\frac{-1}{2\tau_j^2}\beta_j' K_j \beta_j\right),$$

where τ_j^2 is a prior variance determining the impact of the prior distribution on the function estimates, while K_j is the prior precision matrix that implements the smoothness/sparseness assumptions. From a frequentist perspective, these prior distributions correspond to penalties of the form $\lambda_j \beta_j' K_j \beta_j$, where $\lambda_j \geq 0$ is the smoothing parameter that determines the impact of the penalty on the penalised likelihood estimate.

3.3. Model components

By using three specific instances of effects from the class of structured additive predictors, we were able to allow for non-linear, spatial and random effects. For non-linear effects, we relied on Bayesian penalised splines (Eilers and Marx, 1996; Lang and Brezger, 2004), where B-spline basis functions are used to approximate the effect of a continuous covariate. By way of a default, we considered cubic B-splines based on 20 equidistant knots, something that yields twice continuously differentiable (i.e., visually smooth) function estimates, and typically induces enough flexibility to capture even severe non-linearity. As a prior distribution, we used a second-order random walk that, from a frequentist perspective, corresponds to penalising second-order differences of parameters for adjacent basis functions.

For the spatial effect, we used a Markov random field prior that assigns a separate regression coefficient to each region from a discrete set of spatial locations. This implies that the basis functions in this case correspond to simple indicator functions for the different regions. To enforce spatial smoothness, we assumed a Gaussian Markov random field prior for the basis coefficients that induce a penalty where differences between effects of spatially adjacent regions were penalised. For random effects, we assumed independent and identically distributed normal regression coefficients.

A more detailed discussion of potential model components and how they fit into the generic framework outlined above, can be found in Fahrmeir et al. (2013).

3.4. Bayesian inference

For Bayesian inference, we based our approach on the distributional regression framework of Klein et al. (2015), who developed a generic Markov chain Monte Carlo simulation approach, in which proposal densities for blocks of regression coefficients are obtained from a locally quadratic approximation of the log full conditionals. The smoothing variances τ^2 are assigned inverse gamma hyperpriors, such that Gibbs sampling steps can be realised due to the conditionally Gaussian conjugate prior for the regression coefficients. This approach is implemented in the BayesX software package (www.bayesx.org), with the BayesX 3.0.2 version being used for the purposes of our analyses (Belitz et al., 2015).

4. Results

4.1. Model specification

The increased flexibility of distributional regression models for fitting complex response structures, such as that observed for the occurrence and burnt area of wildfires, comes at the price of increased model-choice challenges. More specifically, for each of the four parameters of the zero-one-inflated beta distribution, it was necessary to decide which covariates should be included in the corresponding predictor and which form the corresponding covariate effect should take (e.g., linear versus non-linear effects for continuous covariates).

Based on prior biological and environmental knowledge, this study focused on three meteorological (temperature, precipitation, relative humidity) and one environmental covariate (altitude), while simultaneously accounting for temporal and spatial variability included via a non-linear effect of time (represented by successive days from 1 to 15 August 2006), as well as a spatial effect defined on the basis of the 315 municipalities comprising the study area. The meteorological covariates in particular influenced the state of hydration of dead fuels, with low relative humidity and

high temperatures providing favourable conditions for the initiation and development of wildfires. Lack of precipitation similarly affects both living and dead fuels and, by extension, the occurrence and burnt area of wildfire development.

Based on this pre-selection of covariates, a number of preliminary models were fitted to determine which covariates were associated with specific distributional parameters of the zero-one-inflated beta distribution. Cubic penalised splines, based on twenty inner knots and a second-order random walk prior, were used for potentially non-linear effects. The spatial effect was split into a Markov random field for the spatially structured effect and an independent identically distributed (i.i.d.) random effect for the spatially unstructured effect. For all variance hyperparameters, inverse gamma priors were used, with both parameters set at 0.001.

Based on the consideration of these preliminary models, the following specification was favoured both by theoretical considerations and the comparison of models via the deviance information criterion (DIC):

$$\eta_i^{\mu} = \beta_0^{\mu} + f_1^{\mu}(hr_i) + \beta_2^{\mu} temp_i + f_3^{\mu}(day_i) + \beta_4^{\mu} pp_i + \beta_5^{\mu} altitud_i + f_{str}^{\mu}(cdconc_i) + f_{unstr}^{\mu}(cdconc_i)$$

$$\eta_i^{\sigma^2} = \beta_0^{\sigma^2} + f_1^{\sigma^2}(hr_i) + \beta_2^{\sigma^2} temp_i + f_3^{\sigma^2}(day_i) + \beta_4^{\sigma^2} pp_i + \beta_5^{\sigma^2} altitud_i$$

$$\eta_i^{\nu} = \beta_0^{\nu} + f_1^{\nu}(hr_i) + f_2^{\nu}(temp_i) + f_3^{\nu}(day_i) + \beta_4^{\nu} pp_i + \beta_5^{\nu}(altitud_i) + f_{str}^{\nu}(cdconc_i) + f_{unstr}^{\nu}(cdconc_i)$$

$$\eta_i^{\tau} = \beta_0^{\tau} + f_1^{\tau}(hr_i) + f_3^{\tau}(day_i) + f_{str}^{\tau}(cdconc_i) + f_{unstr}^{\tau}(cdconc_i)$$

For this model, 12,000 iterations of the Markov chain Monte Carlo simulation algorithm were performed, with 2000 iterations being discarded as burn-in and every 10th iteration stored to reduce autocorrelation from the Markov chain. As a consequence, all inferences were based on samples of size 1000 drawn from the posterior predictive distribution of our model.

4.2. Estimated covariate effects

To interpret effects in a distributional regression model, it is necessary to overcome the challenge posed by the common type of *ceteris paribus* analysis, in which changes in one covariate can be related to changes in the expectation of the response while all other covariates are held constant. This change in interpretation results from the fact that covariates generally have an impact on all distributional parameters. However, the advantage of the general distributional regression framework for the zero-one-inflated beta distribution is that it allows one to evaluate effects on many interesting quantities linked to the response distribution and not just effects on the mean. The following measures were considered:

- **The expected burnt area caused by a potential wildfire** (the posterior mean expectation), i.e., the expectation of the zero-one-inflated beta distribution, which is based on all four model parameters and accounts for the probability of no damage (represented by the zero probability), partial damage (represented by the beta part of the distribution) and complete damage (represented by the one probability). As a consequence, the expected burnt area represents an overall risk assessment for specific covariate combinations.
- **The expected burnt area, given fire occurrence** (the posterior mean conditional expectation), i.e., the expected amount of damage, given that there has been a fire at a specific location. This allows for a conditional assessment of wildfire burnt area given fire occurrence, whereas the unconditional expectation represents a mixture of occurrence probability and burnt area.
- **The variability of burnt area, given fire occurrence** (the posterior mean conditional variance), i.e., the variance in the amount of damage, given the occurrence of a fire. This can be interpreted as a measure of risk which quantifies uncertainty in predicting expected burnt area, given fire occurrence. If two locations in an expected burnt area are close together but differ considerably in terms of variability, the location with the higher variance may be considered more risky and this, in turn, can provide additional insights for managing fire risks.
- **The probability of occurrence** (the posterior mean probability of occurrence), i.e., the probability of observing either partial or complete damage at a given location.

Figs. 2–5 depict these quantities for the given covariates. More specifically, the quantities are shown to vary over the range of the respective covariates, while all other covariates are held constant at their averages. Fig. 6 shows the predicted values in comparison with those observed for the burnt area and the probability of occurrence.

This leads to the following conclusions:

- In the case of precipitation, the largest expected burnt area was identified for values of around 4–5 l m^{-2}, i.e., relatively low values. Here it surprised that the burnt area expected for precipitation levels is larger than for dryer levels and this is probably a result of the two opposing effects on the conditional expectation (increasing with precipitation, Fig. 3) and the occurrence probability (decreasing with precipitation, Fig. 5). Taken together, this result in the mode at an intermediate value. For larger amounts of rainfall, the expected burnt area was small. Similarly, the probability of occurrence decreased to zero monotonically with increasing amounts of rainfall and quickly approached zero. Surprisingly, both conditional expectation and conditional variance rose with increasing precipitation and quickly approached their most extreme value of 1. This is likely to be an artefact induced by the linear specification for precipitation in our model (which, however, was favoured in our initial mode comparisons and therefore seems to provide an adequate fit for the majority of the data points).
- In the case of temperature, it is observed that it does not seem to have a great influence on the probability of occurrence and behavior of the wildfires. The origin provoked of practically all of them causes the ambient temperature to become a necessary condition, but not sufficient for the occurrence of the wildfires. In this way, once the average daily temperature reaches about 25–30 °C, an increase does not significantly increase the probability of occurrence of wildfires. As can be seen in the width of the confidence intervals, the enormous variability in the thermometric extremes of the analyzed period is remarkable. In summary, there was a weakly increasing effect on all four parameters of interest, with increasing temperature. While this increasing effect is very plausible, the overall explanatory power of temperature would seem to be small.
- In the case of all four parameters of interest, the highest values were achieved either at a relative humidity of around 30–40% or at a relative humidity of around 80%. For the expected burnt area, there was a clear peak at 30–40% relative humidity. This fits very well with the observation that most fires occurred close to the coast, where low humidity values of under 50% are common. A similar effect structure was found for the probability of occurrence. For the conditional expectation of burnt area, the peaks at 30–40% and at 80% relative humidity were of roughly the same size, while the conditional variance was largest at 80%

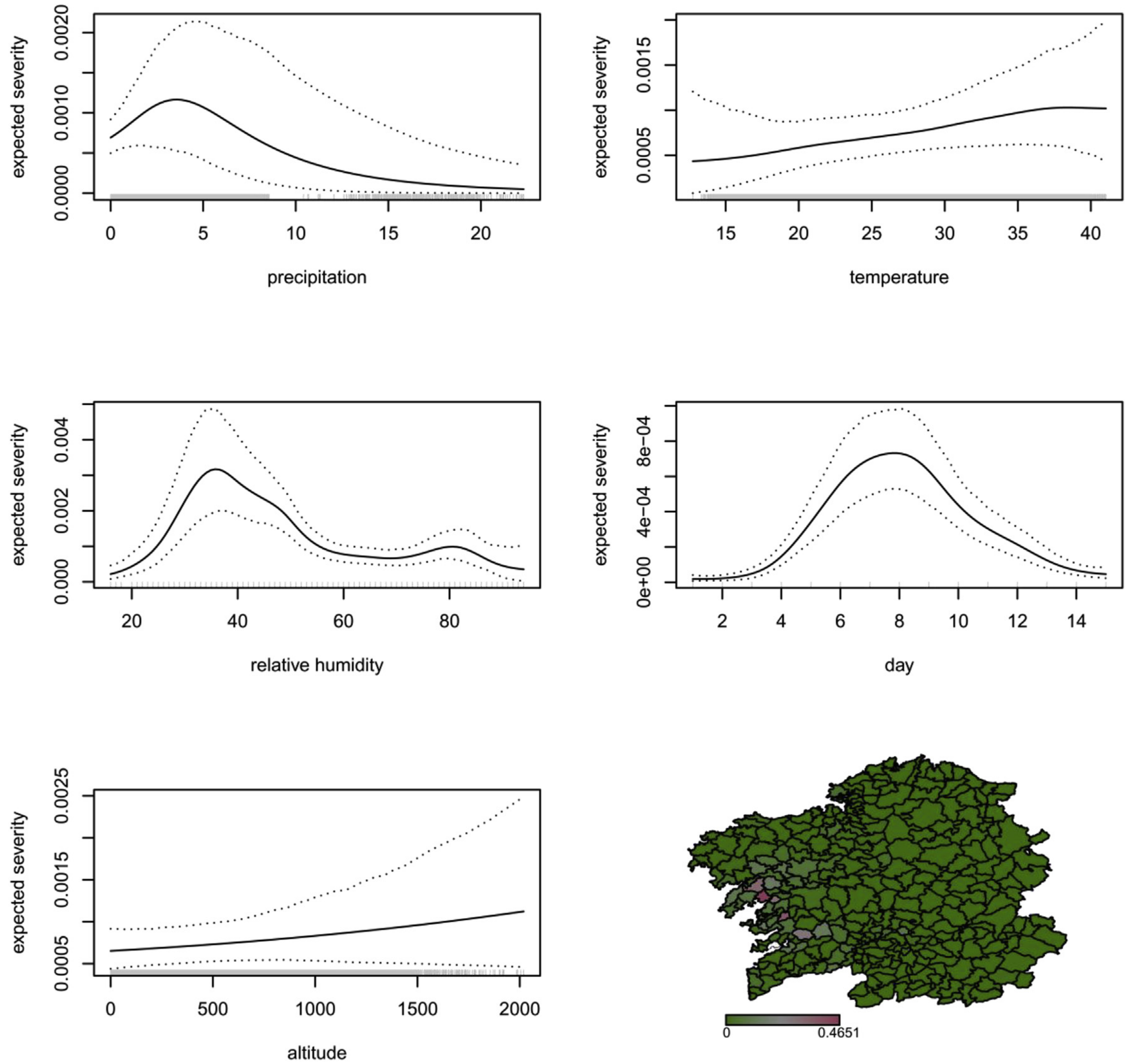

Fig. 2. Posterior mean expectation showing the overall burnt area of wildfire risk varying with respect to precipitation (top left), temperature (top right), relative humidity (centre left), day (centre right), altitude (bottom left) and spatial covariate (bottom right), together with the pointwise 95% credible intervals in the zero-one-inflated beta model.

relative humidity. The latter finding hints at an increased risk at higher humidity values if a fire is actually observed.

- The temporal effect indicated considerable variability over the study period, with a clear peak for all quantities at around the 8th day of the two week period. For the conditional variance, there was a second peak at 12 days. Note that the temporal effect captures all kinds of unobserved heterogeneity that is left after adjusting for the included covariates and is therefore a surrogated of unobserved effects.
- The effect of altitude was not too strong overall but there seemed to be some positive association with the conditionally expected burnt area. There is an increase in variability from 1000 m of altitude. This situation contrasts with that found in other studies of this same research team (Ríos-Pena et al., 2017), where, as the altitude increased, the risk of wildfires decreased. The methodology presented here seems to be more effective in integrating this variable into the predictive model.
- The spatial effect reflected an increased risk towards the coastal areas, a finding that fits well with the clustering of fire events along this strip. Despite allowing for covariates in our model specification, there is obviously still some remaining spatial variation that is captured by the spatial effects.

5. Discussion

Wildfires are a major environmental, economic and social threat. In Mediterranean Europe, they have become the main problem for environmental authorities (Oliveira et al., 2012). In Galicia, they have a remarkable impact on certain areas of the Autonomous Region (*comunidad autónoma*) and pose a challenge for the future of the region. The development of new methodologies, especially those which are evidence-based, allows for more efficient organisation and planning of firefighting, and an ensuingly smaller burnt area and lower risk of loss of life.

This study adopted and tested a new approach to the task of simultaneously studying the occurrence and burnt area of wildfires within the framework of structured additive distributional regression, illustrated by reference to data for Galicia. As shown in Fig. 6, our model seems to fit both the probability of occurrence and the burnt area reasonably well. Structured additive models with non-

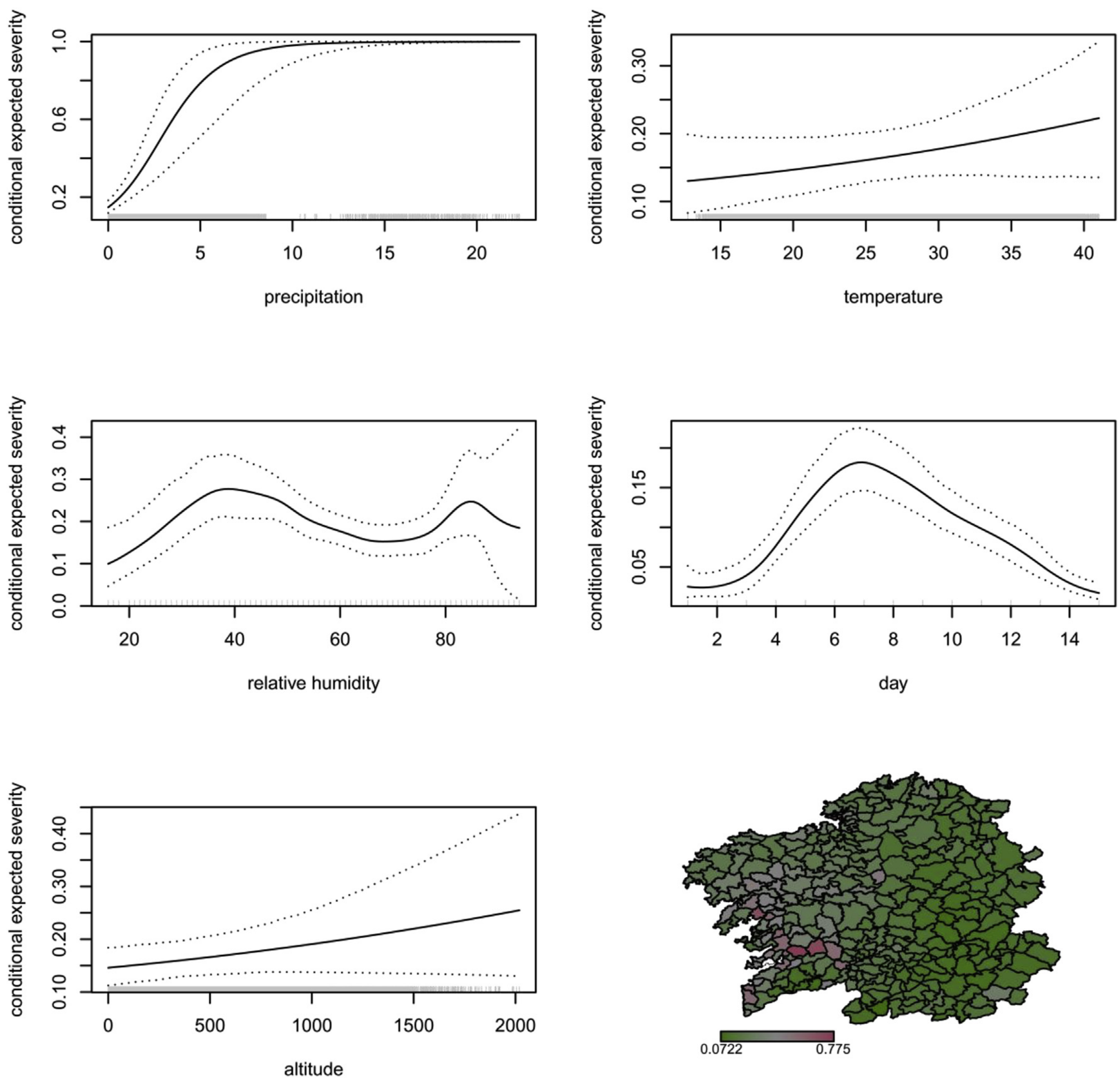

Fig. 3. Posterior mean conditional expectation showing the overall burnt area of wildfire risk varying with respect to precipitation (top left), temperature (top right), relative humidity (centre left), day (centre right), altitude (bottom left) and spatial covariate (bottom right), together with pointwise 95% credible intervals in the zero-one-inflated beta model.

linear interaction effects, which consider spatial information in the form of a location variable, can accelerate the development of fire-behavior models and prove most useful for drawing up fire-prevention and firefighting plans (Ríos-Pena et al., 2017). Their principal advantage lies in their flexibility, in that they can include spatial and temporal covariates along with the potentially non-linear effects of other continuous covariates, i.e., non-linear covariate effects can be combined with the possibility of correcting for spatial autocorrelation through the inclusion of spatial effects (Fahrmeir et al., 2013). In addition the zero-one-inflated beta distribution incorporates the beta distribution with degenerate distributions for the purpose of modelling extreme values. The measures evaluated through estimated covariate effects — namely, expected burnt area of the wildfire, expected burnt area given the occurrence of a wildfire, variability of burnt area given the occurrence of the wildfire, and probability of occurrence of a wildfire at a particular location - showed fire-risk scenarios for the covariates studied.

As indicated by Krasovskii et al. (2016), fire regimes are determined by climate, vegetation and direct human influence. In our case, we have focused on the climatological and physiographic variables. Among the numerous contributions of approaches to wildfire—climate relationships, note the recent works (Littell et al., 2016; Nunes et al., 2016; Turco et al., 2017; Keeley and Syphard, 2017). As expected, spatial and temporal weather conditions influenced the quantity and characteristics of wildfires. Trigo et al. (2013, 2016) established that large summertime wildfire activity was strongly related to previous climatic conditions in the north-western sector of the Iberian Peninsula. Russo et al. (2017) in that the periods of previous drought are a necessary condition for the presence of fires. Our data are centered of the specific days of the arson fires and in which the conditions were adequate for fires spread. In the same way as DaCamara et al. (2014) use generalised Pareto models to improve the general pattern of the Canadian Fire Weather Index (FWI) including among others (temperature, and 24-h cumulated precipitation) one of the most important factors shaping behavior of the arson fires.

If we analyze relative humidity, as it did Bedia et al. (2017) through multivariate spatial models next to ROC curves are proposed. Their results show how the combination of low relative

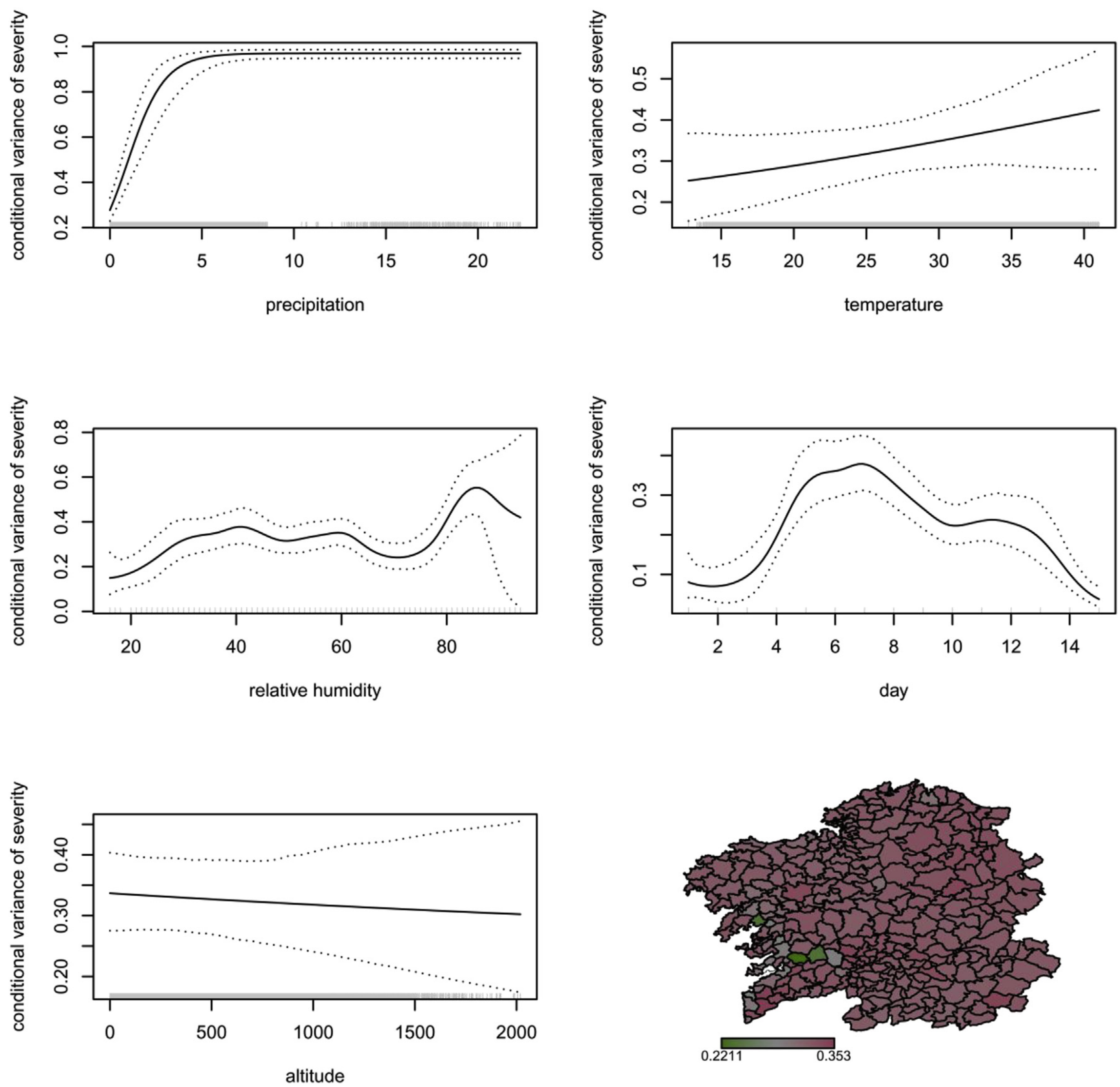

Fig. 4. Posterior mean conditional variance showing wildfire risk varying with respect to precipitation (top left), temperature (top right), relative humidity (centre left), day (centre right), altitude (bottom left) and spatial covariate (bottom right), together with pointwise 95% credible intervals in the zero-one-inflated beta model.

humidity and high air temperature appears in the large fires in the Mediterranean region, which coincides with our results.

In the case of the August 2006 fires in Galicia, these conditions were ideal for fire spread. On analysing and comparing the specific variables of the model, it seems only logical that an increase in temperature would make it easier for wildfires to occur. In the case of Aragon (Spain), Molina-Terrén and Cardil (2016) observed that when the air temperature was higher than 20 °C at 850 hPa, the frequency and burnt area of fires increased. Our results, in terms of average and variance, show that temperature exerts a positive influence on the percentage of burnt area. Oliveira et al. (2014) explained the importance of temperature for wildfire occurrence in Mediterranean Europe, and we found that it was also a key factor for wildfire ignition and behavior. In Greece, Paschalidou and Kassomenos (2016) showed that forest fires and burnt forest area were produced in some very specific weather conditions, similar to our work.

Regarding altitude, our results show that it does not present a significant importance for ignition and fire behavior. It seems an interesting debate because other authors have analyzed the influence of altitude on fires in the Iberian Peninsula. The model proposed by Botequim et al. (2013) to predict annual wildfire risk in “pure and even-aged” *eucalyptus* stands in Portugal, showed altitude as having no significance, coinciding with our results. In a larger-scale study of all wildfires in Portugal from 1987 to 2004, however, Marques et al. (2011) reported that the higher the altitude and distance to roads, the higher the proportion of burnt area. In a nearby area, in contrast, Ordóñez et al. (2012) found altitude to be a negative factor in their generalised linear and generalised least squares models. The explanation given by these authors for this surprising result was that altitude plays a negative role in the probability of a fire being ignited by lightning. Authors like (Rodrigues et al., 2016) conclude that the more mountainous areas to make up for the greater risk of fires with the lowest human activity. It is remarkable the relation between the elevation and the density of roads and its influence on the behavior of the fire. Yang et al. (2007) and Narayaraj and Wimberly (2012) have developed models to measure their influence. In Galicia, Chas Amil et al. (2013) indicate that the density of roads is correlated with fires, but in turn conditioned by population density. Balsa-Barreiro and Hermosilla (2013) concluded in an analysis of the same fires we studied, that the proximity of roads and inhabited areas caused the fires to

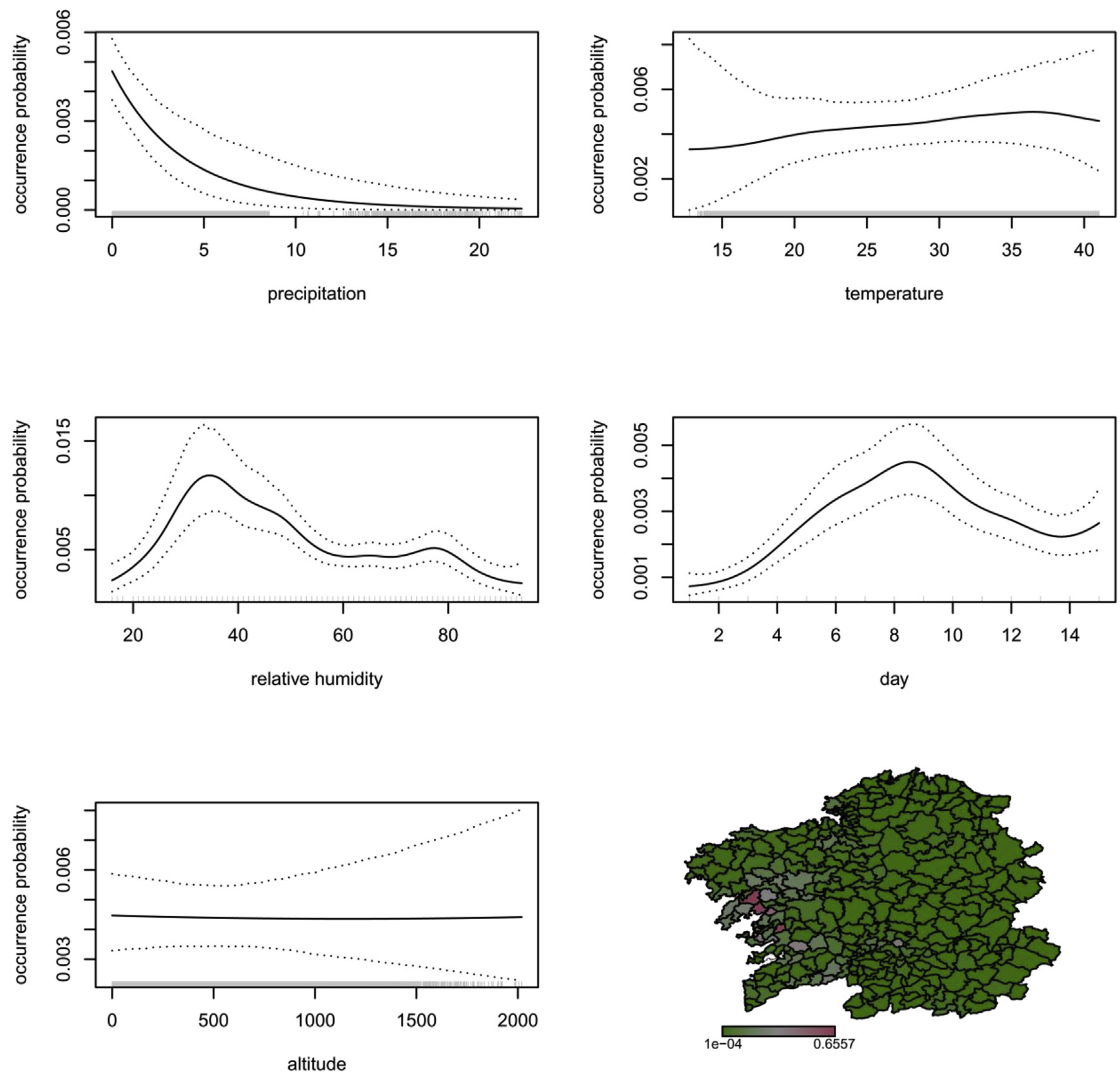

Fig. 5. Posterior mean probability of occurrence showing wildfire risk varying with respect to precipitation (top left), temperature (top right), relative humidity (centre left), day (centre right), altitude (bottom left) and spatial covariate (bottom right), together with pointwise 95% credible intervals in the zero-one-inflated beta model.

increase in size. At the end, the researchers do not reach a consensus on the relationship between altitude and fire, possibly due to the interactions of other variables such as land use, population density and roads.

With respect to the variables with non-linear effects, several authors have incorporated relative humidity (RH) in their indices, and in the Canadian Forest Fire Weather Index in particular (Camia et al., 2008; Dimitrakopoulos et al., 2011; Bedia et al., 2013). Padilla and Vega-García (2011) analyzed the variables that accounted for the fires in each of the 53 regions into which they divided Spain, and concluded that RH provided a good explanation for the fires in the interior of the country but not those in the Galician Region. Pereira et al. (2013) found that the worse wildfire situation in Portugal was due to a combination of wind and low RH. Our results are in line with this finding, in that a RH of around 60% or lower was shown to be accompanied by more and larger fires.

Our analysis of spatial effects showed that fires in August 2006 in Galicia were not distributed randomly but were instead concentrated in certain municipalities in the south-western coastal area, without incidence in the north-eastern region. In the same sense, Chas-Amil et al. (2015) noted that, while the southern and eastern areas acted as "hot spots" with a clustering of parishes with high fire counts, the north-eastern areas acted as "cold spots" with a clustering of parishes with a low number of wildfires.

Other papers of this team and in co-authorship with other researchers have recently used different methodologies to study the distribution of fires in Galicia. Fuentes-Santos et al. (2013) analyzed the spatial distribution of ignition points in the Fonsagrada-Ancares forest district with Ripley's inhomogeneous K-function. Their results indicate that there was spatial dependence between fires within an interaction radius of 4 km for the whole period (1991–2008) and within an interaction radius of 3 km in 1992. Subsequently, Fuentes-Santos et al. (2015) used smooth bootstrap kernel intensity estimation for inhomogeneous spatial point processes to analyze the entire region in 2006, and concluded that the spatial distribution of arson and natural wildfires was different. Boubeta et al. (2015) applied Poisson mixed models with areas included as random effects to 63 forest areas constituting the basic structure of the fight against wildfires in Galicia, during the summer of 2007. The results showed significant differences between forest areas across the period of analysis, similar to our results.

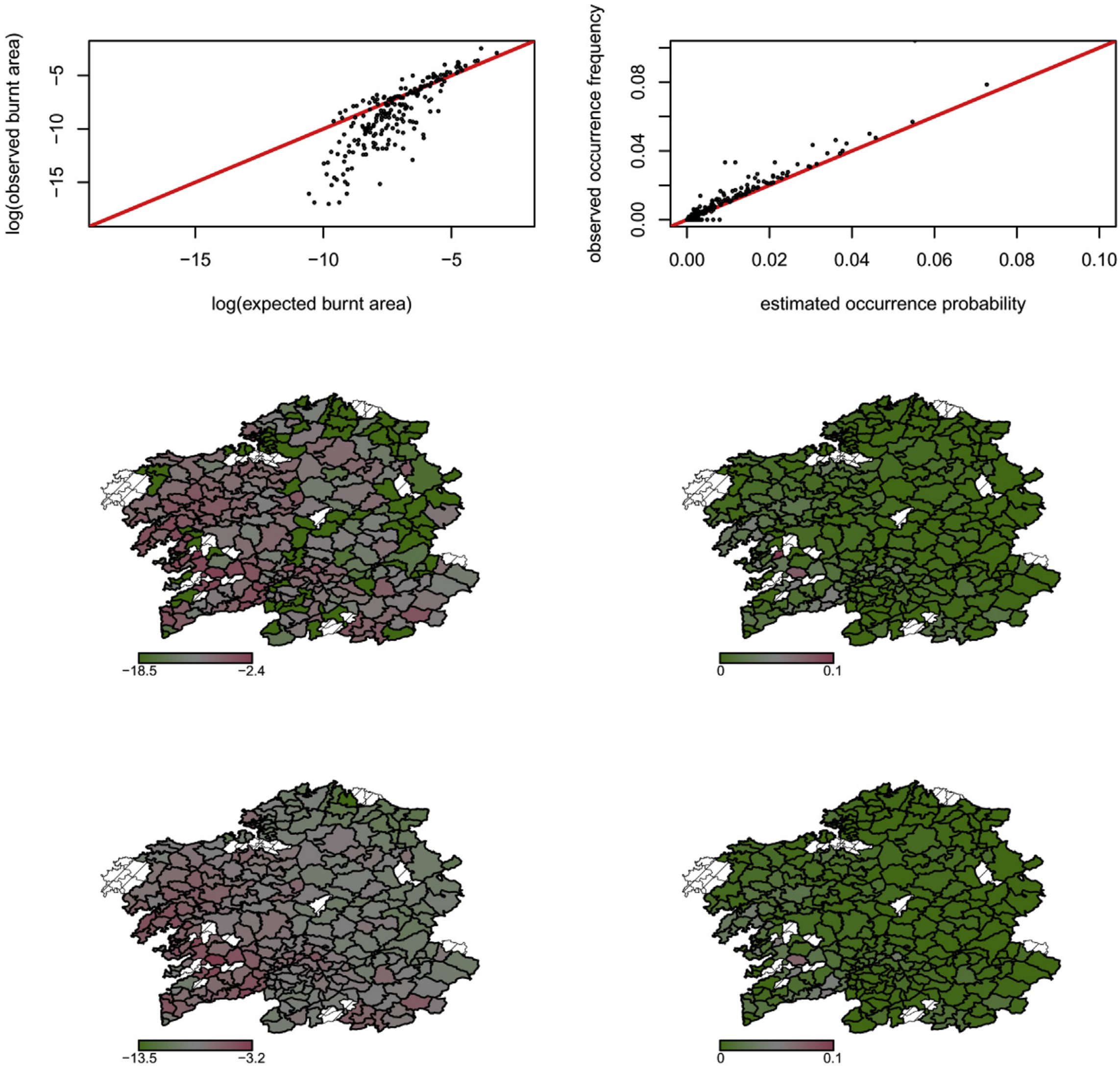

Fig. 6. Representation of the predicted values versus observed values for the burnt area of the wildfires. The top row shows plots of log (burnt area) (observed vs. estimated) and the probability of occurrence (observed vs. estimated). The second row then shows the spatial distribution of the observed log (burnt area) and probability of occurrence while the third row shows the corresponding estimated quantities. All values have been averaged across time as well as all pixels belonging to a certain region.

Lastly, having already developed a sound approach to studying the probability of occurrence of wildfires in space and time by using STAR models for the same data (Ríos-Pena et al., 2015, 2017), we now present a refined and highly reliable method of study and understanding two simultaneous phenomena, i.e., the occurrence of wildfires in a specific place at a given time and the burnt area of the wildfire as measured by the percentage of the area burnt in that place.

6. Conclusions

Since the last third of the 20th century, there has been an increase in the number of forest fires affecting Mediterranean Europe, with repercussions on social, environmental and economic activities in particular. Accordingly, forest service offices and researchers have been developing models capable of predicting fire occurrence zones. Most of the models developed to date have made separate predictions for the number and behavior of fires. In view of the results yielded by this study, STAR models based on the zero-one-inflated beta distribution can be said to afford a simple and satisfactory way of predicting both the probability of the occurrence of a wildfire and, at the same time, the burnt area associated with such a fire. Hence, we suggest that this statistical tool is a novel, methodological contribution to the study of wildfires.

We validated the novel modelling approach on data from Galicia (NW of Spain) in August 2006, when in a rather extreme scenario a total of 83,000 ha were affected by wildfires. This in fact places the model in its operational context. Now and in the future, regional policy makers affected by wildfires, need models of support for extreme situations with simultaneous concentrations of fires in time and space. They need to know the number of events and the surface that can be affected in each climatological situation.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

T. Kneib and C. Cadarso-Suárez would like to thank the Spanish

Ministry of Innovation & Competitiveness for support from the MTM2014-52975-C2-1-R project; and M. Marey-Pérez and C. Cadarso-Suárez would like to thank the Galician Regional Authority (*Xunta de Galicia*) for supporting research of the INBIOEST network (2014-PG141) and Competitive Research Reference Group (PRO-ePLA CN2012/015). The work of N. Klein and T. Kneib was supported by the German Research Foundation (DFG) via the research project KN 922/4-1/2.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.envsoft.2018.03.008.

References

Balsa-Barreiro, J., Hermosilla, T., 2013. Socio-geographic analysis of wildland fires: causes of the 2006 wildfires in Galicia (Spain). Forest Systems 22, 497–509.
Bedia, J., Golding, N., Casanueva, A., Iturbide, M., Buontempo, C., Gutiérrez, J., 2017. Seasonal predictions of fire weather index: paving the way for their operational applicability in Mediterranean Europe. Climate Services. https://doi.org/10.1016/j.cliser.2017.04.001.
Bedia, J., Herrera, S., San Martín, D., Gutíerrez, J., Koutsias, N., 2013. Robust projections of fire weather index in the Mediterranean using statistical downscaling. Climate Change 120, 229–247.
Belitz, C., Brezger, A., Klein, N., Kneib, T., Lang, S., Umlauf, N., 2015. BayesX – software for Bayesian inference in structured additive regression models. Version 3.0.2. Available from: http://www.bayesx.org.
Botequim, B., Garcia-Gonzalo, J., Marques, S., Ricardo, A., Borges, J.G., Tomé, M., Oliveira, M.M., 2013. Developing wildfire risk probability models for Eucalyptus globulus stands in Portugal. iFor. Biogeosci. For. 6, 217–227.
Boubeta, M., Lombardía, M.J., Marey-Perez, M.F., Morales, D., 2015. Prediction of forest fires occurrences with area-level Poisson mixed models. J. Environ. Manag. 154, 151–158.
Brillinger, D.R., Preisler, H.K., Benoit, J.W., 2006. Probabilistic risk assessment for wildfires. Environmetrics 17, 623–633.
Camia, A., Amatulli, G., San Miguel-Ayanz, J., 2008. Past and Future Trends of Forest Fire Danger in Europe. Technical Report EUR 23427 EN - 2008. Institute for Environment and Sustainability. JRC, European Comission, Ispra.
Catry, F.X., Rego, F.C., Bação, F.F., Moreira, F., 2009. Modeling and mapping wildfire ignition risk in Portugal. Int. J. Wildland Fire 18, 921–931.
Chas-Amil, M.L., Prestemon, J.P., McClean, C.J., Touza, J., 2015. Human-ignited wildfire patterns and responses to policy shifts. Appl. Geogr. 56, 164–176.
Chas-Amil, M.L., Touza, J., García-Martínez, E., 2013. Forest fires in the wildland–urban interface: a spatial analysis of forest fragmentation and human impacts. Appl. Geogr. 43, 127–137.
Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M.P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J.R., Herrera, M.A., Zamora, R., 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. Ecol. Model. 221, 46–58.
Conedera, M., Torriani, D., Neff, C., Ricotta, C., Bajocco, S., Pezzatti, G.B., 2011. Using Monte Carlo simulations to estimate relative fire ignition danger in a low-to-medium fire-prone region. For. Ecol. Manag. 261, 2179–2187.
Curt, T., Borgniet, L., Bouillon, Ch, 2013. Wildfire frequency varies with the size and shape of fuel types in southeastern France: implications for environmental management. J. Environ. Manag. 117, 150–161.
DaCamara, C.C., Calado, T.J., Ermida, S.L., Trigo, I.F., Amraoui, M., Turkman, K.F., 2014. Calibration of the Fire Weather Index over Mediterranean Europe based on fire activity retrieved from MSG satellite imagery. Int. J. Wildland Fire 23 (7), 945–958.
Dickson, B.G., Prather, J.W., Xu, Y., Hampton, H.M., Aumack, E.N., Sisk, T.D., 2006. Mapping the probability of large fire occurrence in northern Arizona, USA. Landsc. Ecol. 2, 747–761.
Dimitrakopoulos, A., Bemmerzouk, A., Mitsopoulos, I., 2011. Evaluation of the Canadian fire weather index system in an eastern Mediterranean environment. Meteorol. Appl. 18, 83–93.
Eilers, P.H.C., Marx, B.D., 1996. Flexible smoothing with B-splines and penalties. Stat. Sci. 11, 89–121.
Fahrmeir, L., Kneib, T., Lang, S., 2004. Penalized structured additive regression for spacetime data: a Bayesian perspective. Statistics Sinica 14, 731–761.
Fahrmeir, L., Kneib, T., Lang, S., Marx, B., 2013. Regression: Models, Methods and Applications. Springer, Heidelberg.
Freeborn, P.H., Cochrane, M.A., Matt Jolly, W., 2015. Relationships between fire danger and the daily number and daily growth of active incidents burning in the northern Rocky Mountains, USA. Int. J. Wildland Fire 24, 900–910.
Fréjaville, T., Curt, T., 2015. Spatiotemporal patterns of changes in fire regime and climate: defining the pyroclimates of south-eastern France (Mediterranean Basin). Climatic Change 129, 239–251.
Fuentes-Santos, I., Gonzalez-Manteiga, W., Mateu, J., 2015. Consistent smooth bootstrap kernel intensity estimation for inhomogeneous spatial Poisson point processes. Scand. J. Stat. 10.1111/sjos.12183.
Fuentes-Santos, I., Marey-Pérez, M.F., González-Manteiga, W., 2013. Forest fire spatial pattern analysis in Galicia (NW Spain). J. Environ. Manag. 128, 30–42.
Ganteaume, A., Camia, A., Jappiot, M., San-Miguel-Ayanz, J., Long-Fournel, M., Lampin, C., 2013. A review of the main driving factors of forest fire ignition over Europe. Environ. Manag. 51 (3), 651–662.
González, J.R., Pukkala, T., 2007. Characterization of forest fires in Catalonia (north-east Spain). Eur. J. For. Res. 126, 421–429.
González-Olabarria, J.R., Brotons, Ll, Gritten, D., Tudela, A., Teres, J.A., 2012. Identifying location and causality of fire ignition hotspots in a Mediterranean region. Int. J. Wildland Fire 21, 905–914.
Guiomar, N., Godinho, S., Fernandes, P.M., Machado, R., Neves, N., Fernandes, J.P., 2015. Wildfire patterns and landscape changes in Mediterranean oak woodlands. Sci. Total Environ. 536, 338–352.
Hernandez, Ch, Drobinski, P., Turquety, S., Dupuy, J.-L., 2015. Size of wildfires in the Euro-Mediterranean region: observations and theoretical analysis. Nat. Hazards Earth Syst. Sci. 15, 1331–1341.
Huld, T.A., Súri, M., Dunlop, E.D., Micale, F., 2006. Estimating average daytime and daily temperature profiles within Europe. Environ. Model. Software 21, 1650–1661.
Keeley, J.E., Syphard, D., 2017. Different historical fire–climate patterns in California. Int. J. Wildland Fire 26, 253–268.
Klein, N., Kneib, T., Lang, S., Sohn, A., 2015. Bayesian structured additive distributional regression with an application to regional income inequality in Germany. Ann. Appl. Stat. 9, 1024–1052.
Koutsias, N., Allgöwer, B., Kalabokidis, K., Mallinis, G., Balatsos, P., Goldammer, J.G., 2015. Fire occurrence zoning from local to global scale in the European Mediterranean basin: implications for multi-scale fire management and policy. iForest. https://doi.org/10.3832/ifor1513-008 [online 2015-11-12].
Krasovskii, A., Khabarov, N., Migliavacca, M., Kraxner, F., Obersteiner, M., 2016. Regional aspects of modelling burned areas in Europe. Int. J. Wildland Fire 25, 811–818.
Lampin-Maillet, C., Jappiot, M., Long, M., Bouillon, Ch, Morge, D., Ferrier, J.P., 2010. Mapping wildland-urban interfaces at large scales integrating housing density and vegetation aggregation for fire prevention in the South of France. J. Environ. Manag. 91, 732–741.
Lang, S., Brezger, A., 2004. Bayesian P-splines. J. Comput. Graph Stat. 13, 183–212.
Lasanta, T., Nadal-Romero, E., Arnáez, J., 2015. Managing abandoned farmland to control the impact of re-vegetation on the environment. The state of the art in Europe. Environ. Sci. Pol. 52, 99–109.
Littell, J.S., Peterson, D.L., Riley, K.L., Liu, Y., Luce, C.H., 2016. A review of the relationships between drought and forest fire in the United States. Global Change Biol. 22 (7), 2353–2369.
Loepfe, L., Rodrigo, A., Lloret, F., 2014. Two thresholds determine climatic control of forest fire size in Europe and northern Africa. Reg. Environ. Change 14, 1395–1404.
Marques, S., Borges, J.G., García-Gonzalo, J., Moreira, F., Carreiras, J.M.B., Oliveira, M.M., Cantarinha, A., Botequim, B., Pereira, J.M.C., 2011. Characterization of wildfires in Portugal. Eur. J. For. Res. 130, 775–784.
Martínez-Fernández, J., Chuvieco, E., Koutsias, N., 2013. Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. Nat. Hazards Earth Syst. Sci. 13, 311–327.
Molina-Terrén, D.M., Cardil, A., 2016. Temperature determining larger wildland fires in NE Spain. Theor. Appl. Climatol. 125 (1–2), 295–302.
Nunes, A.N., Lourenço, L., Castro-Meira, A.C., 2016. Exploring spatial patterns and drivers of forest fires in Portugal (1980–2014). Sci. Total Environ. 573, 1190–1202.
Nunez, M., Calhoum, E.A., 1986. A note on air temperature lapse rates on Mount Wellington, Tasmania. Pap. Proc. R. Soc. Tasman. 120, 11–15 (Department of Geography, of Tasmania, Hobart, Australia, and Department of Geography, Newcastle University, Newcastle, Temperature lapse rates are derived from field measurements taken on Mt Wellington over one year).
Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A., Pereira, J.M.C., 2012. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple regression and random forest. For. Ecol. Manag. 275, 117–129.
Oliveira, S., Pereira, J.M.C., San-Miguel-Ayanz, J., Lourenço, L., 2014. Exploring the spatial patterns of fire density in southern Europe using geographically weighted regression. Appl. Geogr. 51, 143–157.
Ordóñez, C., Saavedra, A., Rodríguez-Pérez, J.R., Castedo-Dorado, F., Covián, E., 2012. Using model-based geostatistics to predict lightning-caused wildfires. Environ. Model. Software 29, 44–50.
Orozco, C.V., Tonini, M., Conedera, M., Kanveski, M., 2012. Cluster recognition in spatial-temporal sequences: the case of forest fires. GeoInformatica 16 (4), 653–673.
Ospina, R., Ferrari, S.L.P., 2010. A general class of zero-or-one inflated beta regression models. Comput. Stat. Data Anal. 56, 1609–1623.
Padilla, M.A., Vega-García, C., 2011. On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain. Int. J. Wildland Fire 20, 46–58.
Paschalidou, A.K., Kassomenos, P.A., 2016. What are the most fire-dangerous atmospheric circulations in the Eastern-Mediterranean?. Analysis of the synoptic wildfire climatology. Sci. Total Environ. 539, 536–545.
Paton, D., Tedim, F., 2012. Wildfire and Community: Facilitating Preparedness and

Resilience. Charles C Thomas Publisher Ltd, Springfield Illinois USA.
Pereira, M.G., Caramelo, L., Vega-Orozco, C., Costa, R., Tonini, M., 2015. Space-time clustering analysis performance of an aggregated dataset: the case of wildfires in Portugal. Environ. Model. Software 72, 239–249.
Pereira, M.G., Calado, T.J., DaCamara, C.C., Calheiros, T., 2013. Effects of regional climate change on rural fires in Portugal. Clim. Res. 57 (3), 187–200.
Plucinski, M.P., McCaw, W.L., Gould, J.S., Wotton, B.M., 2014. Predicting the number of daily human-caused bushfires to assist suppression planning in south-west Western Australia. Int. J. Wildland Fire 23, 520–531.
Reid, C.E., Jerrett, M., Petersen, M.L., Pfister, G.G., Morefield, P.E., Tager, I.B., Raffuse, S.M., Balmes, J.R., 2015. Spatiotemporal prediction of fine particulate Matter during the 2008 northern California wildfires using Machine learning. Environ. Sci. Technol. 49, 3887–3896.
Rigby, R.A., Stasinopoulos, D.M., 2005. Generalized additive models for location, scale and shape. Apploed Statistics 54, 507–554.
Ríos-Pena, L., Cadarso-Suárez, C., Kneib, T., Marey-Pérez, M.F., 2015. Applying binary structured additive regression (STAR) for predicting wildfire in Galicia, Spain. Procedia Environmental Sciences 27, 123–126.
Ríos-Pena, L., Kneib, T., Cadarso-Suárez, C., Marey-Pérez, M., 2017. Predicting the occurrence of wildfires with binary structured additive regression models. J. Environ. Manag. 187, 154–165.
Rodrigues, M., Jiménez, A., de la Riva, J., 2016. Analysis of recent spatial-temporal evolution of human driving factors of wildfires in Spain. Nat. Hazards 84, 2049–2070.
Rodrigues, M., de la Riva, J., Fotheringham, S., 2014. Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Appl. Geogr. 48, 52–63.
Ruffault, J., Moron, V., Trigo, R.M., Curt, T., 2016. Daily synoptic conditions associated with large fire occurrence in Mediterranean France: evidence for a wind-driven fire regime. Int. J. Climatol. https://doi.org/10.1002/joc.4680.
Ruiz-Mirazo, J., Martínez-Fernández, J., Vega-García, C., 2012. Pastoral wildfires in the Mediterranean: understanding their linkages to land cover patterns in managed landscapes. J. Environ. Manag. 98, 43–50.
Russo, A., Gouveia, C.M., Páscoa, P., DaCamara, C.C., Sousa, P.M., Trigo, R.M., 2017. Assessing the role of drought events on wildfires in the Iberian Peninsula. Agric. For. Meteorol. 237, 50–59.
San-Roman-Sanz, A., Fernandez, C., Mouillot, F., Ferrat, L., Istria, D., Pasqualini, V., 2013. Longterm forest dynamics and land-use abandonment in the Mediterranean mountains, Corsica, France. Ecol. Soc. 18, 38.
Thompson, M.P., Calkin, D.E., 2011. Uncertainty and risk in wildland fire management: a review. J. Environ. Manag. 92, 1895–1909.
Trigo, R.M., Sousa, P.M., Pereira, M.G., Rasilla, D., Gouveia, C.M., 2016. Modelling wildfire activity in Iberia with different atmospheric circulation weather types. Int. J. Climatol. 36, 2761–2778.
Trigo, R.M., Sousa, P.M., Pereira, M.G., Rasilla, D., Gouveia, C.M., 2013. Modelling wildfire activity in Iberia with different atmospheric circulation weather types. Int. J. Climatol. 10.1002/joc.3749.
Turco, M., Von Hardenberg, J., AghaKouchak, A., Llasat, M.C., Provenzale, A., Trigo, R.M., 2017. On the key role of droughts in the dynamics of summer fires in Mediterranean Europe. Nature – Scientific Reports 7, 81. https://doi.org/10.1038/s41598-017-00116-9.
Verde, J.C., Zêzere, J.L., 2010. Assessment and validation of wildfire susceptibility and hazard in Portugal. Nat. Hazards Earth Syst. Sci. 10, 485–497.
Viedma, O., Moity, N., Moreno, J.M., 2015. Changes in landscape fire-hazard during the second half of the 20th century: agriculture abandonment and the changing role of driving factors. Agric. Ecosyst. Environ. 207, 126–140.
Weise, D.R., Wotton, B.M., 2010. Wildland-urban interface fire behaviour and fire modelling in live fuels. Int. J. Wildland Fire 19 (2), 253–271.
Yang, J., Hong, S.H., Shifley, S.R., Gustafson, E.J., 2007. Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands. For. Sci. 53, 1–15.
Yiannakoulias, N., Kielasinska, E., 2015. The effect of temperature on arson incidence in Toronto, Ontario, Canada. Int. J. Biometeorol. 1–11.

Further reading

Comas, C., Costafreda-Aumedes, S., Vega-García, C., 2014. Characterizing configurations of fire ignition points through spatiotemporal point processes. Nat. Hazards Earth Syst. Sci. 2, 2891–2911.
Ganteaume, A., Long-Fourmel, M., 2015. Driving factors of fire density can spatially vary at the local scale in south-eastern France. Int. J. Wildland Fire 24, 650–664.
González, J.R., Palahí, M., Tasobares, A., Pukkala, T., 2006. A fire probability model for forest stands in Catalonia (north-east Spain). Ann. For. Sci. 63, 169–176.
Khbarov, N., Krasovskii, A., Obersteiner, M., Swart, R., Dosio, A., San-Miguel-Ayanz, J., Durrant, T., Camia, A., Migliavacca, M., 2014. Forest fires and adaptation options in Europe. Reg. Environ. Change 16, 21–30.
Lawrence, M.G., 2005. The relationship between relative humidity and the dewpoint temperature in moist air: a simple conversion and applications. Bull. Am. Meteorol. Soc. 86, 225–233. https://doi.org/10.1175/BAMS-86-2-225.